

AD 643557

ON A SPECIAL CLASS OF CONSTRAINED GENERALIZED MEDIAN PROBLEMS

A. CHARNES

M. J. L. KIRBY
University of Chicago

and

W. M. RAIKE

TECHNICAL REPORT 66-107

NOVEMBER 1966



CLEARINGHOUSE FOR FEDERAL SCIENTIFIC AND TECHNICAL INFORMATION		
Hardcopy	Microfiche	
\$ 3.00	\$.65	24 pp 20
1 ARCHIVE COPY		

**ON A SPECIAL CLASS OF CONSTRAINED
GENERALIZED MEDIAN PROBLEMS**

A. CHARNES

M. J. L. KIRBY
University of Chicago

and

W. M. RAIKE

TECHNICAL REPORT 66-107

NOVEMBER 1966



**Information-Processing And Control
SYSTEMS LABORATORY**

NORTHWESTERN UNIVERSITY

EVANSTON, ILLINOIS

60201

Reproduction in whole or in part is permitted for any purpose of the United States Government.

ACCESSION for	
CFSTI	WHITE SECTION <input checked="" type="checkbox"/>
CDC	BUFF SECTION <input type="checkbox"/>
UNANNOUNCED	<i>Per statement on Doc</i>
IDENTIFICATION	
Y <i>Im</i>	
DISTRIBUTION/AVAILABILITY CODES	
DIST.	AVAIL. and/or SPECIAL
<i>[Handwritten mark]</i>	

The research described in this report was jointly supported by the U. S. Army, Navy, and Air Force in a Joint Service Electronics Program under Office of Naval Research Contract Number N00014-66-C0020-A03 (Identification Number NR 373-502/3-14-66 Electronics Branch) and the U. S. Army under U. S. Army Research Office — Durham Contract Number DA-31-124-ARO-D-322.

The distribution of this document is unlimited. Qualified requesters may obtain copies of this report from the Defense Documentation Center.

TECHNICAL REPORT 66-107

ON A SPECIAL CLASS OF CONSTRAINED GENERALIZED MEDIAN PROBLEMS

A. Charnes

Mathematics Department, Economics Department,
Engineering Sciences Department, and, Industrial
Engineering and Management Sciences Department

M. J. L. Kirby

University of Chicago

W. M. Raike

Industrial Engineering and Management Sciences Department

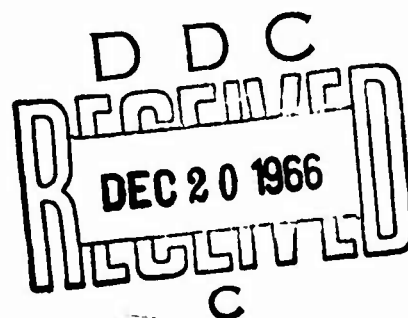
The research described in this report was jointly supported by the U. S. Army, Navy, and Air Force in a Joint Service Electronics Program under Office of Naval Research Contract Number N00014-66-C0020-A03 (Identification Number NR 373-502/3-14-66 Electronics Branch) and the U. S. Army under U. S. Army Research Office-Durham Contract Number DA-31-124-ARO-D-322.

November 1966

Information-Processing And Control

SYSTEMS LABORATORY

NORTHWESTERN UNIVERSITY
EVANSTON, ILLINOIS
60201



CONTENTS

Abstract	ii
1. Introduction	1 - 5
2. Mathematical Development	5 - 12
3. The Increment and the Radial Directional Derivative of $h(X)$13 - 14
4. Extensions15 - 16
References17
Document Control Data- R&D (DD Form 1473)	18 - 19

ABSTRACT

By means of elementary properties of the absolute value function, important properties of a special class of "constrained generalized median" problems (and eventually, the most general class, vide Charnes, Cooper, Thompson) such as existence of solutions, gradient and incremented formulae, linear programming and probabilistic interpretations are obtained for all classes of joint distribution functions for which the problems make sense. Results of A. C. Williams and R. Wets obtained by involved arguments and sophisticated constructs appear, when corrected, as special instances of some of the above results but devoid of the interrelations and interpretations herein adduced.

1. Introduction

Recently the SIAM Journal has published two papers [1, 8] which present elaborate classifications and sophisticated constructions for two classes of problems in probabilistic programming. Neither the authors of these papers, nor their referees, appear to have recognized that the problem in [1] and the problems in [8] (for example, the "complete" problem) are very special instances of the class of "constrained generalized median" problems first discussed by Charnes, Cooper and Thompson in 1961.^{1/} This canonical "median" formulation is not a matter of a choice of nomenclature. Its technical advantage in many problems in probabilistic programming derives from bringing into immediate focus the relevance of the absolute value function.

As we shall demonstrate below by means of the absolute value function we easily (1) obtain, characterize and markedly extend the essential results of [1], (2) interpret these results in linear programming and probabilistic terms, (3) develop gradients and directional derivatives and interpret probabilistically incremental formulae in all generality, (4) place these results in the framework of chance-constrained programming, and (5) reduce all cases to investigation of the behavior of the objective function along a ray.

Before undertaking these developments and in order not to interrupt them some important assertions and constructions in [1] require correction. The problem in [1] is to choose a vector $X \geq 0$ which maximizes the expected value of the function $p^T X + \sum_{i=1}^m g_i(X, b)$.

^{1/} See reference [4]

where

$$(1) \quad g_i(X, b) = \begin{cases} (b_i - a^i X) \gamma_i & \text{if } b_i - a^i X \leq 0 \\ (b_i - a^i X) \delta_i & \text{if } b_i - a^i X \geq 0, \end{cases}$$

and $\gamma_i \geq \delta_i$.

Here $a^i, i=1 \dots m$, is the i^{th} row of a given $m \times n$ matrix A ,

γ_i and $\delta_i, i=1 \dots m$, are given constants,

p^T is a given $1 \times n$ vector,

and $b_i, i=1, \dots, m$, are random variables whose marginal distributions are known.

Thus the problem is

$$(2) \quad \begin{aligned} & \max E(p^T X + \sum_{i=1}^m g_i(X, b)) \\ & \text{subject to } X \geq 0, \end{aligned}$$

where we compute the expectation using the joint distribution of the random variables b_1, \dots, b_m .

In (2), it is assumed that the vector X is to be selected before any observations are made on the random variables $b_i, i=1, \dots, m$. Thus X is not permitted to be a function of these random variables, but rather it must be a deterministic vector. Such a vector of decision rules is called a zero order decision rule in the customary terminology of chance-constrained programming.^{1/} Thus, (2) is a chance-constrained programming problem in which the chance constraints are of the particularly simple form $P(X \geq 0) \geq 1$, in which the function whose expectation we wish to maximize is a piecewise linear function of the vector of decision rules, and in which we seek the optimal zero order rule.

^{1/}For results on zero order rules see references [2, 3, 4, 5, 6].

In the mathematical development presented in section 2, we first show that (2) can be converted into a constrained generalized median problem of the type discussed in [4]. This transformation enables us to use standard linear programming methods in obtaining our results. It also permits us to indicate how these results can be extended in two important directions: one is the case in which $g_1(X, b)$ is a piecewise linear function of b_1, \dots, b_m rather than only b_1 as it is in (1); the other is the case in which X is subjected to other linear constraints, say $DX \geq d$, in addition to the constraints $X \geq 0$. Neither of these extensions are discussed in [1], and the techniques used in [1] cannot be readily extended to include these cases.

The main results in [1] are a pair of theorems which give necessary and sufficient conditions for the existence of an optimal solution to (2). We will obtain most of these conditions in a much more direct manner than was used in [1]. In particular, our proofs will require only well-known theorems in linear programming and some elementary inequalities on the absolute value, and will not require concepts such as Kakutani's fixed point theorem which were used in [1]. Neither will we require the elaborate specifications and restrictions of the classes of probability distributions as used in [1].

One result obtained in [1] characterizes the situation in which the objective function of (2) is bounded from above, but the supremum of the function is not attained for a finite X . Rather than deriving this particular result, we characterize this situation in a constructive way by showing that it can be related to the behavior of the objective function along a particular ray and by obtaining an explicit expression for its limiting value.

First, however, we must correct some assertions in [1] about the admissible class of random variables. In [1] it is assumed that F_1 , the cumulative distribution function of b_1 , has at most a finite number of discontinuities in each finite interval. In our development in section 2 it will be clear that such an assumption is never required; when differentiability is not the main focus. We can, for example, admit random variables whose distribution function has probability mass at all the irrational points in an interval (or on the real line).

It is also assumed in [1] that $F_1(\cdot), i=1, \dots, m$, satisfies the following conditions:

$$3(a) \quad \lim_{z_1 \rightarrow -\infty} z_1 F_1(z_1) = 0,$$

and

$$3(b) \quad \lim_{z_1 \rightarrow +\infty} z_1 [1 - F_1(z_1)] = 0.$$

A remark is then made [1, p. 930] to the effect that 3(a) and 3(b) are "sufficient for (but slightly stronger than) the existence of the first moment of $F_1(\cdot)$."

In actual fact, as we prove in theorem 2, 3(a) and 3(b) are necessary for the existence of the first moment of $F_1(\cdot)$, but they are not sufficient. This can be seen by the following example:

$$\text{Let } F(z) = \begin{cases} 1, & z \geq -e \\ \frac{e}{|z| \ln|z|}, & z < -e \end{cases}$$

Then

$$\lim_{z \rightarrow -\infty} zF(z) = \lim_{z \rightarrow -\infty} \frac{-e}{\ln|z|} = 0$$

and

$$\lim_{z \rightarrow +\infty} z[1 - F(z)] = 0, \text{ as } F(z) = 1 \text{ for } z \geq -e$$

Hence $F(z)$ satisfies conditions 3(a) and 3(b). However, $F(z)$ does not have a

first moment since

$$\begin{aligned}
 \int_{-\infty}^{+\infty} z dF(z) &= e \int_{-\infty}^{-e} z \left(\frac{1+\ln(-z)}{z^2 [\ln(-z)]^2} \right) dz \\
 &= -e \int_e^{\infty} \frac{1+\ln(z)}{z [\ln(z)]^2} dz \\
 &= -e \int_e^{\infty} \frac{1}{z [\ln(z)]^2} dz + e \left[\ln(\ln(e)) - \lim_{t \rightarrow +\infty} [\ln(\ln(t))] \right]
 \end{aligned}$$

$= -\infty$, since the integral term is negative and $\lim_{t \rightarrow +\infty} [\ln(\ln(t))] = \infty$. Thus $E(z)$ does not exist.

Conditions 3(a) and 3(b) are used repeatedly in [1] to guarantee that the integration by parts formula

$$\int_{-\infty}^x F_1(z) dz = xF_1(x) - \lim_{z \rightarrow -\infty} zF_1(z) - \int_{-\infty}^x z dF_1(z)$$

is valid. But the above example shows that conditions 3(a) and 3(b) are not sufficient to guarantee the validity of this integration formula since both integrals diverge. On the other hand, the existence of $E(z)$ is certainly a sufficient condition for the formula to be meaningful.

2. Mathematical Development

We begin by introducing some notation. Let Y be an n -vector.

Then $|Y^T| = (|y_1|, \dots, |y_1|, \dots, |y_n|)$

and

$$E|Y^T| = (E|y_1|, \dots, E|y_1|, \dots, E|y_n|).$$

We will use $||Y||$ to denote the norm of Y .

From the definition of $g_i(X, b)$ in (1) we have that

$$g_i(X, b) = \frac{\delta_i}{2} (|b_i - a^i X| + (b_i - a^i X)) + \frac{\gamma_i}{2} ((b_i - a^i X) - |b_i - a^i X|).$$

Hence

$$E(g_i(X, b)) = \left(\frac{\delta_i + \gamma_i}{2}\right) E(b_i) - \left(\frac{\delta_i + \gamma_i}{2}\right) a^i X + \left(\frac{\delta_i - \gamma_i}{2}\right) E|b_i - a^i X|$$

and

$$\begin{aligned} E\left(p^T X + \sum_{i=1}^m g_i(X, b)\right) &= p^T X + \sum_{i=1}^m E(g_i(X, b)) = \sum_{i=1}^m \left(\frac{\delta_i + \gamma_i}{2}\right) E(b_i) \\ &+ \left(p^T - \sum_{i=1}^m \left(\frac{\delta_i + \gamma_i}{2}\right) a^i\right) X + \sum_{i=1}^m \left(\frac{\delta_i - \gamma_i}{2}\right) E|b_i - a^i X|. \end{aligned}$$

Thus, dropping the constant term $\sum_{i=1}^m \left(\frac{\delta_i + \gamma_i}{2}\right) E(b_i)$ from the objective function, (2) can be written as

$$(4) \quad \min h(X) = c^T X + \alpha^T E|b - AX|$$

subject to $X \geq 0$

where
$$c_j = \sum_{i=1}^m \left(\frac{\delta_i + \gamma_i}{2}\right) a_{ij} - p_j, \quad j = 1, \dots, n,$$

and
$$\alpha_i = \frac{\gamma_i - \delta_i}{2}, \quad i = 1, \dots, m.$$

Also $\alpha_i \geq 0$ as we are given that $\gamma_i \geq \delta_i$.

(4) is a constrained generalized median problem of the type discussed in [4].

The constraints are of the particularly simple form $X \geq 0$.

We first note that

Theorem 1: $h(X)$ is finite for some $X \geq 0$, in which case

$h(X)$ is finite for all $X \geq 0$, if and only if $E|b_i| < \infty$ for all i such that $\alpha_i > 0$ (i.e. $\gamma_i > \delta_i$).

Proof: Using the well-known and elementary inequalities on the absolute value we have that

$$\pm |AX| \mp |b| \leq \left| |AX| - |b| \right| \leq |b - AX| \leq |b| + |AX|$$

and so applying the expected value operator (which is order-preserving) to both sides of the inequalities we get

$$(c^T X + \alpha^T |AX|) + \alpha^T E|b| \leq h(X) \leq (c^T X + \alpha^T |AX|) + \alpha^T E|b|.$$

This establishes the sufficiency and necessity of the stated conditions:

Theorem 2: A necessary condition for $E|b_1| < \infty$ is that

$$\lim_{t \rightarrow \infty} tP\{|b_1| \geq t\} = 0.$$

Proof:

$$\text{Let } \chi_t = \begin{cases} 1, & |b_1| \leq t \\ 0, & |b_1| > t \end{cases}$$

$$\text{Then } E|b_1| = E(|b_1|\chi_t) + E(|b_1|(1 - \chi_t)).$$

But by definition of the Lebesgue-Stieltjes (or Radon) integral,

$$E|b_1| = \lim_{t \rightarrow \infty} E(|b_1|\chi_t) \text{ when } E|b_1| < \infty.$$

$$\text{Thus } E(|b_1|(1 - \chi_t)) \rightarrow 0 \text{ as } t \rightarrow \infty.$$

$$\text{But } E(|b_1|(1 - \chi_t)) = \int_t^\infty |b_1| dF_1(b_1) \geq tP\{|b_1| \geq t\} \geq 0 \text{ for } t \geq 0.$$

$$\text{So when } E|b_1| < \infty \text{ we have } tP\{|b_1| \geq t\} \rightarrow 0 \text{ as } t \rightarrow \infty.$$

To see that the condition

$$\lim_{t \rightarrow +\infty} tP\{|b_1| \geq t\} = 0$$

is equivalent to 3(a) and 3(b), note that

$$P\{|b_1| \geq t\} = P(b_1 \geq t) + P(b_1 \leq -t) = 1 - F_1(t) + F_1(-t).$$

$$\text{Hence we have } 0 = \lim_{t \rightarrow +\infty} (t[1 - F_1(t)] + tF_1(-t)),$$

and since $F_1(\cdot)$ is a nonnegative function, 3(a) and 3(b) immediately result.

Theorem 3: $h(X)$ is bounded from below for all $X \geq 0$ if and only if there exists a w satisfying

$$\begin{aligned} |w| &\leq \alpha \\ w^T A + c^T &\geq 0. \end{aligned}$$

Proof: By virtue of the inequality

$$c^T X + \alpha^T |AX| - \alpha^T E|b| \leq h(X) \leq c^T X + \alpha^T |AX| + \alpha^T E|b|$$

obtained in the proof of theorem 1, it is enough to show that $c^T X + \alpha^T |AX|$ is bounded from below for all $X \geq 0$ if and only if there exists a w satisfying $|w| \leq \alpha$ and $w^T A + c^T \geq 0$.

To do this, we first note that the problem

$$(5) \quad \begin{aligned} \min \quad & c^T X + \alpha^T |AX| \\ \text{subject to} \quad & X \geq 0 \end{aligned}$$

can be rewritten in the following linear programming form:

$$\begin{aligned} \min \quad & c^T X + \alpha^T (Y^+ + Y^-) \\ \text{subject to} \quad & Y^+ - Y^- - AX = 0 \\ & X, Y^+, Y^- \geq 0. \end{aligned}$$

The dual to this problem is

$$\begin{aligned} \max \quad & w^T \cdot 0 \\ \text{subject to} \quad & -w^T A \leq c^T \\ & w^T \leq \alpha^T \\ & -w^T \leq \alpha^T \end{aligned}$$

or

$$\begin{aligned} \max \quad & w^T \cdot 0 \\ \text{subject to} \quad & w^T A + c^T \geq 0 \\ & |w| \leq \alpha. \end{aligned}$$

Since (5) is consistent (the origin is feasible), the extended dual theorem of linear programming (see [7], vol. I, p. 190) states that the objective function of (5) is finite if and only if there is a feasible solution to its dual, i.e. (6). Thus the theorem is proved.

To convert the result of theorem 3 to the corresponding result in [1], we use our definitions of c^T and α . Since $\alpha_1 = \frac{\gamma_1 - \delta_1}{2}$, we see that $|w_1| \leq \alpha$ means $\frac{\delta_1 - \gamma_1}{2} \leq w_1 \leq \frac{\gamma_1 - \delta_1}{2}$. If we put $\pi_1 = w_1 + \left(\frac{\gamma_1 + \delta_1}{2}\right)$, then the condition $|w_1| \leq \alpha$ becomes $\delta_1 \leq \pi_1 \leq \gamma_1$. Similarly, since $c_j = \sum_{i=1}^m \left(\frac{\delta_i + \gamma_i}{2}\right) a_{ij} - p_j$, we get that $w^T A + c^T = \pi^T A - p^T$, so $w^T A + c^T \geq 0$ becomes $\pi^T A \geq p^T$. Thus we see that theorem 3 is analogous to the result in [1] which says that the optimal value of the objective function to (2) is finite if and only if there exists an m -vector π satisfying

$$\pi^T A \geq p^T$$

$$\text{and} \quad \delta \leq \pi \leq \gamma.$$

We now turn to an elucidation of what is termed in [1] the "insoluble^{1/} finite" case. This is the situation we mentioned in the introduction in which $h(X)$ is bounded from below but its infimum is not attained for a finite X , i.e. $h(X) > \inf_{X \geq 0} h(X)$ for all $X \geq 0$.

Theorem 4: As $t \rightarrow \infty$, $E|b_1 - tq| \sim \begin{cases} tq - E(b_1), & q > 0 \\ E(b_1) - tq, & q < 0 \\ E|b_1|, & q = 0, \end{cases}$

where we use the symbol \sim to mean that the difference between the two quantities tends to zero as $t \rightarrow \infty$.

^{1/}This terminology is peculiar since not only is it not insoluble but we here explicitly write down the optimal value of the objective function for this case. This is done in the corollary to theorem 4.

Proof:

Define the function $\chi_{tq} = \chi_{tq}(b_i) = \begin{cases} 1, & b_i < tq \\ 0, & b_i \geq tq \end{cases}$.

$$\begin{aligned} \text{Then } E|b_i - tq| &= E[(b_i - tq)(1 - \chi_{tq})] + E[(tq - b_i)\chi_{tq}] \\ &= \left(E(b_i) - 2E(b_i \chi_{tq}) \right) + tq[P(b_i < tq) - P(b_i \geq tq)]. \end{aligned}$$

$$\text{But } E(b_i \chi_{tq}) \sim \begin{cases} 0, & q < 0 \text{ as } t \rightarrow \infty, \\ E(b_i), & q > 0 \end{cases}$$

and the second term on the right approaches tq if $q > 0$ and $-tq$ if $q < 0$. Thus the theorem is proved.

Now fix X and define $I^0 = \{i : a^i X = 0\}$. Let $b_i^1 = -b_i \operatorname{sgn}(a^i X)$ and $b_i^2 = \begin{cases} b_i, & i \in I^0 \\ 0, & i \notin I^0 \end{cases}$. Then we have

Corollary: As $t \rightarrow \infty$, $h(tX) \sim t[c^T X + \alpha^T |AX|] + \alpha^T [E(b^1) + E|b^2|]$.

Proof:

$$\begin{aligned} h(tX) &= tc^T X + \alpha^T E|b - tAX| \sim tc^T X + \sum_{i \in I^0} \alpha_i E|b_i| \\ &\quad + \sum_{i \notin I^0} \alpha_i \operatorname{sgn}(a^i X) [ta^i X - E(b_i)] \\ &= t[c^T X + \alpha^T |AX|] + \alpha^T [E(b^1) + E|b^2|]. \end{aligned}$$

Theorem 5: Suppose that $h(X_n) \rightarrow v_0$ for some sequence $\{X_n\}$ with $X_n \geq 0$, $\|X_n\| \rightarrow \infty$ as $n \rightarrow \infty$, where $v_0 = \inf_{X \geq 0} h(X)$. Then $c^T \bar{Y} + \alpha^T |A\bar{Y}| = 0$ for some $\bar{Y} \geq 0$ with $\|\bar{Y}\| = 1$.

Proof:

Since the set $S_1 = \{X : \|X\| = 1, X \geq 0\}$ is compact, the set

$\left\{ \frac{X_n}{\|X_n\|} \right\}$ has a limit point \bar{Y} in S_1 . Moreover, there exists some

subsequence, which we denote by $\{Y_n\}$, such that $Y_n \rightarrow \bar{Y}$ as $n \rightarrow \infty$.

Now

(7)

$$\frac{1}{\|X_n\|} (c^T X_n + \alpha^T |AX_n|) = c^T Y_n + \alpha^T |AY_n| \rightarrow c^T \bar{Y} + \alpha^T |\bar{A}\bar{Y}| \text{ as } n \rightarrow \infty.$$

However, $c^T X_n + \alpha^T |AX_n|$ is bounded from below since $h(X_n) > v_0$.

This follows from the inequality $(c^T X + \alpha^T |AX|) - \alpha^T E|b| \leq h(X)$

$\leq (c^T X + \alpha^T |AX|) + \alpha^T E|b|$. Thus, in particular, $c^T X_n + \alpha^T |AX_n|$

remains bounded while $\|X_n\| \rightarrow \infty$. Hence, from (7), we conclude

that the left hand side goes to zero while $c^T \bar{Y} + \alpha^T |\bar{A}\bar{Y}| \geq 0$.

Thus the theorem is proved.

Theorem 6:

A necessary and sufficient condition that $\lim_{t \rightarrow \infty} h(tX) = \inf_{t \geq 0} h(tX)$ is that $c^T X + \alpha^T |AX| = 0$.

Proof: Assume $\|X\| = 1$ since the conclusion is trivial for $X=0$ and

$c^T X + \alpha^T |AX|$ is positive homogeneous. Suppose $c^T X + \alpha^T |AX| \neq 0$. The

Corollary to Theorem 4 guarantees that $\lim_{t \rightarrow \infty} h(tX) = \lim_{t \rightarrow \infty} t [c^T X + \alpha^T |AX|] +$

$\alpha^T E|b| + \alpha^T E|b|^2$, so that the facts that $c^T X + \alpha^T |AX| \neq 0$ and that $E|b_i| < \infty$

for all i prove that $\lim_{t \rightarrow \infty} h(tX)$ exists and is finite. Now suppose that

there is a t_0 such that $\inf_{t \geq 0} h(tX) = h(t_0 X)$. Since $h(tX)$ is a convex func-

tion of t , $h(\frac{1}{2}t_0 X + \frac{1}{2}tX) \leq \frac{1}{2} [h(t_0 X) + h(tX)]$ for all t . But $\lim_{t \rightarrow \infty} h(tX) =$

$\lim_{t \rightarrow \infty} h(\frac{1}{2}t_0 X + \frac{1}{2}tX) = v_0$, so that $v_0 \leq \frac{1}{2} h(t_0 X) + \frac{1}{2} v_0$, or $v_0 \leq h(t_0 X)$. Since

$h(t_0 X) \leq v_0$ by definition of t_0 , we have shown $\inf_{t \geq 0} h(tX) = h(t_0 X) = v_0 =$

$\lim_{t \rightarrow \infty} h(tX)$ when t_0 exists; if there is no such t_0 than a fortiori

$\inf_{t \geq 0} h(tX) = \lim_{t \rightarrow \infty} h(tX)$.

To prove the converse, note that if $X_n = t_n X$ for some sequence having

$t_n \rightarrow \infty$, then $\frac{X_n}{\|X_n\|} = X$ for all n . Theorem 5 then immediately yields the desired results.

We cast the following evident observation as a theorem because of its practical utility:

Theorem 7: $\inf_{X \geq 0} h(X) = \inf_{\|X\|=1} \inf_{t \geq 0} h(tX).$

The usefulness of this observation rests on Theorem 6. In applications it may happen that there is only one \bar{X} with $\|\bar{X}\| = 1$ for which $c^T \bar{X} + \alpha^T |A\bar{X}| = 0$. In cases such as this, Theorem 7 shows that one can approximate the value of v_0 as closely as desired at some finite X by considering only ray minima for rays tX with X sufficiently near \bar{X} .

However, it should be expressly noted that these "knife-edge" infinite ray cases can never appear as solutions to practical problems. Rather they exhibit an inadequacy of realistic formulation of the model. We develop their properties here only for completeness of analysis.

We now turn to the development of incremental formulae and their interpretation in probabilistic terms. It should be noted that, again, elementary properties of absolute values quickly and naturally lead to results in all generality.

3. The Increment and the Radial Directional Derivative of $h(X)$.

We now give expressions in terms of the probabilities of the various b versus AX events for the increment of $h(X)$ (i.e., $h(X + \xi) - h(X)$) and for the directional derivative $\frac{d}{dt} h(tX)$ whenever the latter exists. The expressions are free of any restrictions as to the character of the distributions of the b .

First, on inspecting figure 1 below, we note that, for $\epsilon > 0$,

$$f(u) = |u - (a + \epsilon)| - |u - a| = \begin{cases} \epsilon & , u < a \\ \epsilon + 2(a - u) & , a \leq u < a + \epsilon \\ -\epsilon & , u \geq a + \epsilon \end{cases}$$

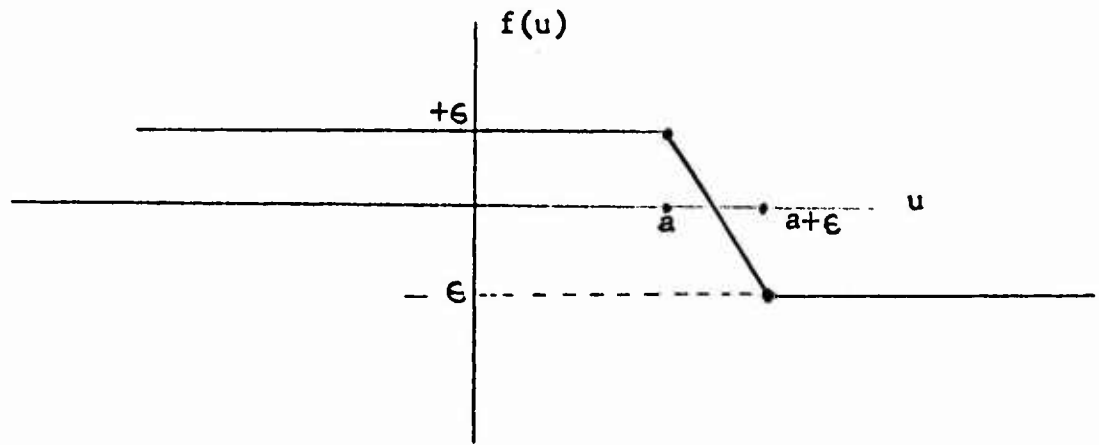


Figure 1

Thus, if u is a random variable with finite expectation,

$$E_u \left(|u - (a + \epsilon)| - |u - a| \right) = \epsilon [P(u < a + \epsilon) - P(u \geq a + \epsilon)] + \left[\epsilon + 2a - 2(a + \epsilon \theta(\epsilon)) \right] P(a \leq u < a + \epsilon)$$

for some $0 \leq \theta(\epsilon) \leq 1$ by the mean value theorem. In other words,

$$(8) \quad E_u \left(|u - (a + \epsilon)| - |u - a| \right) = \epsilon [P(u < a + \epsilon) - P(u \geq a + \epsilon) - 2\theta(\epsilon)P(a \leq u < a + \epsilon)].$$

But

$$h(X+\xi) - h(X) = c^T \xi + \alpha^T E \left(|b - AX - A\xi| - |b - AX| \right),$$

so if we use (8) we get, using $P(b \geq AX)$ to denote the vector whose components are $P(b_i \geq a^i x)$,

$$(9) \quad h(X+\xi) - h(X) = c^T \xi + \alpha^T D(A\xi) \left[P(b < A(X+\xi)) - P(b \geq A(X+\xi)) \right. \\ \left. - 2D(\theta(|A\xi|)) P(AX \leq b < A(X+\xi)) \right]$$

where $D(y)$ denotes the diagonal matrix whose diagonal consists of the components of the vector y and where $\theta(|A\xi|) = (\theta_1(|a^1 \xi|), \dots, \theta_m(|a^m \xi|))$ with $0 \leq \theta_i(|a^i \xi|) \leq 1$. Thus we have an expression for the increment of $h(X)$, and thereby an expression for any possible directional variation, in terms of the probabilities of various events. Clearly then, any conditions on the probability distributions involved in b which guarantee that $P(AX \leq b < A(X+\xi)) \rightarrow 0$ as $\|\xi\| \rightarrow 0$ will guarantee the existence of a gradient of $h(X)$ at X .

We now specialize (9) as follows: In (9) replace X by tX and ξ by ϵX .

Then we get

$$h(t+\epsilon)X - h(tX) = c^T(\epsilon X) + \alpha^T D(A\epsilon X) \left[P(b < A(t+\epsilon)X) - P(b \geq A(t+\epsilon)X) \right. \\ \left. - 2D(\theta(|A\epsilon X|)) P(AtX \leq b < A(t+\epsilon)X) \right] = \epsilon \left\{ c^T X + \alpha^T D(AX) \left[P(b < (t+\epsilon)AX) \right. \right. \\ \left. \left. - P(b \geq (t+\epsilon)AX) - 2D(\theta(|AX|)) P(tAX \leq b < (t+\epsilon)AX) \right] \right\}.$$

Then as $t \rightarrow \infty$ we get, for any fixed $\epsilon > 0$,

$$(10) \quad \frac{h((t+\epsilon)X) - h(tX)}{\epsilon} \sim c^T X + \alpha^T |AX|.$$

Thus, if the left hand side of (10) has a limit as $\epsilon \rightarrow 0$, we see that the directional derivative $\frac{d}{dt} h(tX)$ is given by

$$(11) \quad \lim_{t \rightarrow \infty} \frac{d}{dt} h(tX) = c^T X + \alpha^T |AX|.$$

It is clear that computational methods can be based on our expressions (9), (10) and (11), but we shall reserve these developments for another occasion.

4. Extensions

The special model treated in [1] is rarely of any real-world or economic-theoretical significance because many constraints cannot be expressed adequately by "linear" penalty functions. As a simple reminder, note that X must be bounded, although of course the specification of the bounds may involve interrelations between groups of the individual variables. We have chosen to develop our results for clarity and ease of comparison in the context of [1]. However, it should be obvious on reflection that practically all the developments of the preceding sections can be extended to the case where there are additional linear inequality constraints on X in (2).

For example, consider the extension of theorem 3 to the case where X must also satisfy $DX \geq d$. Then our proof of theorem 3 requires only the obvious modification of the dual problem. Thus we get

Theorem 8: If the constraints $DX \geq d$, $X \geq 0$ are consistent, the optimal value of the objective function for the problem

$$\begin{aligned} \min \quad & c^T X + \alpha^T E |b - AX| \\ \text{subject to} \quad & DX \geq d \\ & X \geq 0 \end{aligned}$$

is finite if and only if there exist vectors w and v satisfying

$$\begin{aligned} -w^T A + v^T D &\leq c^T \\ |w| &\leq \alpha \\ v &\geq 0 \end{aligned}$$

A much more significant extension of the general model follows from observing that the linear constraints $DX \geq d$ can be considered to be the deterministic equivalent constraints for chance constraints $P(DX \leq b) \geq \beta$. This follows from

the fact that since X is a zero order rule, the chance constraints can be replaced by the equivalent set of deterministic constraints $DX \leq F^{-1}(1-\beta)$, where $F^{-1}(1-\beta)$ is the vector of $F_i^{-1}(1-\beta_i)$, the $(1-\beta_i)$ fractile point of the marginal distribution of b_i .

It is also evident that modifying the objective function to the form

$$\min c^T X + E|Gb - AX|,$$

where G is a given $m \times m$ matrix, does not affect the proof of our theorems.

Thus our results hold for the very general class of constrained generalized median models discussed in [4].

Finally, it should be noted that the so-called "complete" problem^{1/} of linear programming under uncertainty is merely another special instance of the constrained generalized median problem, and as such falls under our and the previous analysis in [4].

To see this we merely observe that the "complete" problem can be written as

$$\begin{aligned} \min \quad & c^T X + E \left(\sum_{i=1}^m g_i(X, b) \right) \\ \text{s.t.} \quad & DX = d \\ & X \geq 0 \end{aligned}$$

$$\text{where } g_i(X, b) = \begin{cases} (b_i - a_i^1 X) \gamma_i & \text{if } b_i - a_i^1 X \leq 0 \\ (b_i - a_i^1 X) \delta_i & \text{if } b_i - a_i^1 X \geq 0. \end{cases}$$

This formulation is obtained by expressing the second stage decision variables in terms of the first stage decision variables X .

^{1/}See [8], page 102.

"Completely slacked" might be more appropriate since 'complete' implies generality rather than the speciality actually involved.

References

- [1] Williams, A. C., "On Stochastic Linear Programming," Journal of the Society for Industrial and Applied Mathematics, Vol. 13, No. 4, December, 1965, pp. 927-40.
- [2] Charnes, A. and W. W. Cooper, "Deterministic Equivalents for Optimizing and Satisficing Under Chance Constraints," Operations Research, Vol. 11, No. 1, January-February, 1963, pp. 18-39.
- [3] Kirby, M. J. L., "Generalized Inverses and Chance-Constrained Programming," Ph.D. Thesis in Applied Mathematics, Northwestern University, Evanston, Illinois, June, 1965.
- [4] Charnes, A., W. W. Cooper, and G. L. Thompson, "Constrained Generalized Medians and Hypermedians as Deterministic Equivalents for Two-Stage Linear Programs Under Uncertainty," Management Science, Vol. 12, No. 1, September, 1965, pp. 83-112. Also O.N.R. Research Memorandum No. 41, Northwestern University, Evanston, Illinois July 6, 1961.
- [5] Ben Israel, A., "On Some Problems of Mathematical Programming," Ph.D. Thesis in Engineering Science, Northwestern University, Evanston, Illinois, June, 1962.
- [6] Charnes, A. and M. Kirby, "Optimal Decision Rules for the n-Period E-Model of Chance-Constrained Programming," Cahiers du Centre d'Etudes de Recherche Operationnelle. January, 1966.
- [7] Charnes, A. and W. W. Cooper, Management Models and Industrial Applications of Linear Programming, Vols. I and II. New York: John Wiley and Sons, Inc., 1961.
- [8] Wets, R. J. B., "Programming Under Uncertainty: The Equivalent Convex Program," Journal of the Society for Industrial and Applied Mathematics, Vol. 14, No. 1, Jan. 1966, pp. 89-105.

Security Classification

DOCUMENT CONTROL DATA - R&D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY (Corporate author)		2a. REPORT SECURITY CLASSIFICATION	
Information- Processing And Control Systems Laboratory Northwestern University, Evanston, Illinois 60201		Unclassified	
		2b. GROUP	
		Not applicable	
3. REPORT TITLE			
ON A SPECIAL CLASS OF CONSTRAINED GENERALIZED MEDIAN PROBLEMS			
4. DESCRIPTIVE NOTES (Type of report and inclusive dates)			
Not applicable			
5. AUTHOR(S) (Last name, first name, initial)			
Charnes, Abraham Kirby, Michael J. L. University of Chicago Raikes, William M.			
6. REPORT DATE		7a. TOTAL NO. OF PAGES	7b. NO. OF REFS
November 1966		22	8
8a. CONTRACT OR GRANT NO.		9a. ORIGINATOR'S REPORT NUMBER(S)	
Contract Number N00014-66-C0020-A03 Contract Number DA-31-124-ARO-D-322 b. PROJECT NO.		66-107	
NR 373-502/3-14-66 Electronics Branch			
Not applicable		9b. OTHER REPORT NO(S) (Any other numbers that may be assigned this report)	
d. Not applicable		Northwestern University Systems Research Memorandum No. 155	
10. AVAILABILITY/LIMITATION NOTICES			
Qualified requesters may obtain copies of this report from the Defense Documentation Center.			
11. SUPPLEMENTARY NOTES		12. SPONSORING MILITARY ACTIVITY	
Not applicable		Office of Naval Research, Washington, D. C. U. S. Army Research Office, Durham, N. C.	
13. ABSTRACT			
<p>By means of elementary properties of the absolute value function, important properties of a special class of "constrained generalized median" problems (and eventually, the most general class, vide Charnes, Cooper, Thompson) such as existence of solutions, gradient and incremented formulae, linear programming and probabilistic interpretations are obtained for all classes of joint distribution functions for which the problems make sense. Results of A. C. Williams and R. Wets obtained by involved arguments and sophisticated constructs appear, when corrected, as special instances of some of the above results but devoid of the interrelations and interpretations herein adduced.</p>			

Unclassified

Security Classification

14. KEY WORDS	LINK A		LINK B		LINK C
	ROLE	WT	ROLE	WT	ROLE
chance-constrained programming linear programming under uncertainty constrained generalized medians zero-order decision rules					

INSTRUCTIONS

1. **ORIGINATING ACTIVITY:** Enter the name and address of the contractor, subcontractor, grantee, Department of Defense activity or other organization (corporate author) issuing the report.

2a. **REPORT SECURITY CLASSIFICATION:** Enter the overall security classification of the report. Indicate whether "Restricted Data" is included. Marking is to be in accordance with appropriate security regulations.

2b. **GROUP:** Automatic downgrading is specified in DoD Directive 5200.10 and Armed Forces Industrial Manual. Enter the group number. Also, when applicable, show that optional markings have been used for Group 3 and Group 4 as authorized.

3. **REPORT TITLE:** Enter the complete report title in all capital letters. Titles in all cases should be unclassified. If a meaningful title cannot be selected without classification, show title classification in all capitals in parenthesis immediately following the title.

4. **DESCRIPTIVE NOTES:** If appropriate, enter the type of report, e.g., interim, progress, summary, annual, or final. Give the inclusive dates when a specific reporting period is covered.

5. **AUTHOR(S):** Enter the name(s) of author(s) as shown on or in the report. Enter last name, first name, middle initial. If military, show rank and branch of service. The name of the principal author is an absolute minimum requirement.

6. **REPORT DATE:** Enter the date of the report as day, month, year; or month, year. If more than one date appears on the report, use date of publication.

7a. **TOTAL NUMBER OF PAGES:** The total page count should follow normal pagination procedures, i.e., enter the number of pages containing information.

7b. **NUMBER OF REFERENCES:** Enter the total number of references cited in the report.

8a. **CONTRACT OR GRANT NUMBER:** If appropriate, enter the applicable number of the contract or grant under which the report was written.

8b, 8c, & 8d. **PROJECT NUMBER:** Enter the appropriate military department identification, such as project number, subproject number, system numbers, task number, etc.

9a. **ORIGINATOR'S REPORT NUMBER(S):** Enter the official report number by which the document will be identified and controlled by the originating activity. This number must be unique to this report.

9b. **OTHER REPORT NUMBER(S):** If the report has been assigned any other report numbers (either by the originator or by the sponsor), also enter this number(s).

10. **AVAILABILITY/LIMITATION NOTICES:** Enter any limitations on further dissemination of the report, other than those

imposed by security classification, using standard statements such as:

- (1) "Qualified requesters may obtain copies of this report from DDC."
- (2) "Foreign announcement and dissemination of this report by DDC is not authorized."
- (3) "U. S. Government agencies may obtain copies of this report directly from DDC. Other qualified DDC users shall request through _____."
- (4) "U. S. military agencies may obtain copies of this report directly from DDC. Other qualified users shall request through _____."
- (5) "All distribution of this report is controlled. Qualified DDC users shall request through _____."

If the report has been furnished to the Office of Technical Services, Department of Commerce, for sale to the public, indicate this fact and enter the price, if known.

11. **SUPPLEMENTARY NOTES:** Use for additional explanatory notes.

12. **SPONSORING MILITARY ACTIVITY:** Enter the name of the departmental project office or laboratory sponsoring (paying for) the research and development. Include address.

13. **ABSTRACT:** Enter an abstract giving a brief and factual summary of the document indicative of the report, even though it may also appear elsewhere in the body of the technical report. If additional space is required, a continuation sheet shall be attached.

It is highly desirable that the abstract of classified reports be unclassified. Each paragraph of the abstract shall end with an indication of the military security classification of the information in the paragraph, represented as (TS), (S), (C), or (U).

There is no limitation on the length of the abstract. However, the suggested length is from 150 to 225 words.

14. **KEY WORDS:** Key words are technically meaningful terms or short phrases that characterize a report and may be used as index entries for cataloging the report. Key words must be selected so that no security classification is required. Identifiers, such as equipment model designation, trade name, military project code name, geographic location, may be used as key words but will be followed by an indication of technical context. The assignment of links, roles, and weights is optional.